JRC Scientific and Technical Reports

Scientific, Technical and Economic Committee for Fisheries (STECF)

Report of the Working Group on review of economic data collected in relation to the DCF, harmonisation of sampling strategies (SGEGA 10-03)

20-24 SEPTEMBER 2010, SALERNO

EDITED BY EVELINA SABATELLA AND JARNO VIRTANEN

EUR 24635 EN - 2010

JRC EUROPEAN COMMISSION

ipsc Institute for the Protection and Security of the Citizen

The mission of the Institute for the Protection and Security of the Citizen (IPSC) is to provide research results and to support EU policy-makers in their effort towards global security and towards protection of European citizens from accidents, deliberate attacks, fraud and illegal actions against EU policies

The Scientific, Technical and Economic Committee for Fisheries (STECF) has been established by the European Commission. The STECF is being consulted at regular intervals on matters pertaining to the conservation and management of living aquatic resources, including biological, economic, environmental, social and technical considerations.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Address: TP 051, 21027 Ispra (VA), Italy
E-mail: stecf-secretariat@jrc.ec.europa.eu
Tel.: 0039 0332 789343
Fax: 0039 0332 789658

https://stecf.jrc.ec.europa.eu/home
http://ipsc.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC 61980

EUR 24635 EN
ISBN 978-92-79-18752-0
ISSN 1831-9424
doi:10.2788/58635

Luxembourg: Publications Office of the European Union

Printed in Italy

TABLE OF CONTENTS

SCIENTIFIC, TECHNICAL AND ECONOMIC COMMITTEE FOR FISHERIES (STECF)

STECF Comments on the report of the Sub Group on Economic Affairs (SGECA 10-03) Review of economic data collected in relation to the DCF, harmonisation of sampling strategies

STECF OPINION EXPRESSED DURING THE PLENARY MEETING (PLEN 10-03) held in Brussels, 8-12 November 2010

Request to STECF

STECF was requested to review the report of the STECF-SGECA-10-03 Working Groups of September 20 - 24, 2010 (Salerno) meeting, evaluate the findings and make any appropriate comments and recommendations.

Background

NP proposals 2011-2013 were submitted by MS under the new guidelines proposed by the STECF-SGRN 09-03 WG and endorsed by STECF. The guidelines take into account the recommendations from the STECF-SGECA-09-02 WG. In particular, the STECF-SGECA-09-02 WG suggested that in order to obtain methodology descriptions of a comparable standard among MS, a methodological report should be included in the national programme proposals. The aim of the methodological report is to provide a clear and detailed description of the data collection methods.

The STECF-SGECA 10-03 WG was asked to evaluate these methods with the specific objectives to compare different approaches and to discuss possible harmonisation of sampling strategies among MS.

In addition, the DCF requires STECF to evaluate the quality of the data collected by Member States (article 7, Council Regulation (EC) 199/2008). The STECF-SGECA 10-03 WG was asked to suggest proper procedures to facilitate STECF's task.

STECF comments and recommendations

STECF notes that the STECF-SGECA 10-03 WG interpreted ToR 2. to refer to "proposed" indicators in National Programmes rather than "achieved" indicators in Annual Reports.

STECF recognises that SGECA 10-03 addressed all their terms of reference and, in addition, provided useful guidelines to assess the quality of methodologies and of estimates.

STECF interprets the phrase "harmonisation of sampling strategies" to mean: making the sampling strategies such that the results of the surveys are comparable.

STECF notes that STECF-SGRN WGs have provided Guidelines to MS on the preparation of National Programmes and Guidelines on the preparation of Annual Reports. STECF agrees with the additional guidelines on statistical techniques for simple and stratified random sampling reported in Annex III of the report of the STECF-SGECA 10-03 WG. The Commission is asked to circulate these guidelines to National Correspondents.

STECF reminds the Commission that the STECF-SGECA 09-02 WG recommended a study to standardise quality reporting and propose methods in the case of non probability sampling survey. Terms of Reference for this study are in the report of the STECF-SGECA 09-02 WG and have

already been endorsed by STECF (July 2009). STECF reiterates its recommendation to conduct the study as soon as possible so that MS can consider the results of the study when presenting quality indicators in the 2011 technical report on activities performed in 2010 (to be delivered by May 2011). STECF proposes that the TOR of this study should also include options to deal with possible bias by non-probability sampling.

STECF proposes to evaluate proposed methods in National Programmes in addition to evaluating compliance with DCF regulations in forthcoming STECF-SGRN WGs. To enable this additional evaluation, it is intended that appropriate criteria will be developed in a Workshop on Statistical issues in 2011. Future STECF-SGRN WGs could also evaluate the quality of elements included in National Annual Reports, as well as whether the required elements have been provided.

STECF notes that the STECF-SGECA 10-03 WG did not provide revised methods for the collecting aquaculture data because of time constraints and lack of appropriate expertise. STECF considers that a review of proposed methods in National Programmes for collecting aquaculture data would be useful and could generate valuable advice to MS.

STECF reviewed the guidelines proposed by the STECF-SGECA 10-03 WG on how MS should collect and present information on quality of the data collected, which include definitions of accuracy indicators to be presented by MS in the Annual Report and the type of presentation for each indicator. STECF notes that, in the proposed guidelines, the use of the term census is not consistent. STECF proposes to use the following definition for census: Census is a type data collection in which every unit in the frame population is approached for data. If there are non-responses when the census is carried out, then the response rate will be below 100%.

STECF notes that there is a table of definitions and guidelines to present accuracy indicators included in the report of the STECF-SGECA 10-03 WG. STECF considers that this table should be revised as suggested below to improve clarity of explanation and then included in Guidelines to be distributed to MS.

Table 1: Definition and presentation of accuracy indicators to be presented by MS in the AR

Type of data collection	**Accuracy indicators**	**Definition and *presentation***
A: Census	Response rate	achieved no.[(1)] / frame population no. *Present as %*
	Coefficient of Variation (CV) only if response rate <70%[(2)]	$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$ where: $\hat{\sigma}(\hat{Y})$ is the estimated standard deviation of $\hat{Y}$ $\hat{Y}$ is the estimated total value per fleet segment of the variable e.g. total energy costs[(3)] *Present as absolute term (e.g. 0.2 rather than 20%)*
B: Probability Sample survey C: Non-Probability Sample survey	Achieved sampling rate	achieved sample no.[(4)] /frame population no. *Present as %*
	Response rate	achieved sample no.[(4)] / planned sample no.[(5)] *Present as %*
	Coefficient of Variation (CV)	$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$ where: $\hat{\sigma}(\hat{Y})$ is the estimated standard deviation of $\hat{Y}$ $\hat{Y}$ is the estimate of the total [(3)] *Present as absolute term (e.g. 0.2 rather than 20%)*

(1) **Achieved no.** is the number of respondents who supplied data in response to the census
(2) **CV** is also required for census which achieves a low response rate (<70%) as this must be treated as if it were a Non-Probability Sample survey
(3) The **estimated total** is the final estimate for each variable and each fleet segment, according to appendix VI of DCF. E.g. estimated total energy costs, estimated total crew costs, per fleet segment
(4) **Achieved sample no.** is the number of respondents that supply data (and not, for instance, the number of questionnaires sent out, or number of companies contacted)
(5) **Planned sample no.** is the number of units to be contacted for the survey (even though you may not expect all of them to respond and supply data)

STECF proposes that Guidelines to MS for presentation of Annual Reports are revised at the next relevant STECF-SGRN WG meeting, taking into account the recommendations in annex IV of the STECF-SGECA 10-03 WG report. Guidelines for the processing and aquaculture sectors should also be revised following the same approach.

STECF notes that, in order to evaluate the quality of data collected in Annual Reports, there should be specific targets for CV, acceptable response rates, etc. STECF agrees with the method suggested by the STECF-SGECA 10-03 WG to apply a step by step procedure in order to define targets for quality evaluation. In the first step (for 2011 Annual Reports), targets for quality assessment will be qualitative, e.g. MS should provide transparent and complete descriptions of methods and analysis of accuracy indicators. The second step is to analyse and evaluate quality indicators in order to help define targets for data quality. This step should ideally be carried out in an *ad hoc* workshop convened by the Commission, rather than an STECF WG, and should include appropriate experts appointed by MS.

STECF notes that the Perpetual Inventory Method (PIM) to estimate fleet capital values, as required by the DCF, creates comparable results on the capital values of fishing fleets among MS. STECF notes that the DCF requires MS to estimate both replacement and historical values, which could create ambiguity when one value has to be chosen for economic analysis. STECF notes that, although valuation at historical prices is usual in company accounts, it cannot be compared with national accounting or other economic statistics that are expressed at prices of a single period. Therefore, STECF **recommends** use of replacement value in macroeconomic analyses (as in the case of AER).

STECF recognizes that there is a wide range of possible input data (price per capacity unit, interest rates, depreciation times and methods) for the Perpetual Inventory Method and **recommends** that further more detailed guidelines be developed to specify which input data is preferred.

STECF agrees with the STECF-SGECA 10-03 WG's recommendation to use the geometric approach to estimating depreciation, as recommended by OECD (OECD manual on measuring capital stocks, 2009).

With regard to the processing sector, STECF **recommends** that MS avoid duplication of sampling effort, according to statistical best practices. Therefore, in cases where official statistics on economic data on the processing sector cannot be used to meet the requirements of the DCF, MS should clearly explain the reason and justify the use of additional surveys.

STECF **recommends** that MS describe definitions and methods used to collect economic data regardless the source used in order to ensure transparency.

STECF **recommends** use of the IVth Council Directive 78/660/EEC as reference in order to standardise definition adopted by MS. In particular, net financial costs are defined in Article 23, item C.13 of the IVth Council Directive 78/660/EEC, ("Interest payable and similar charges, with a separate indication of those concerning affiliated undertakings"), while net extraordinary costs are defined in Article 23, item 17, ("Extraordinary charges").

Noting the Commission request to provide a definition of processing industry and activities, STECF notes that the Statistical classification of economic activities in the European Community (NACE Rev. 2) includes a detailed description of activity 10.20 "Processing and preserving of fish, crustaceans and molluscs" and **recommends** MS use this description when identifying the population that should be surveyed. However, STECF also notes that the current DCF states that the population shall refer to enterprises whose main activity is defined according to the

EUROSTAT definition under NACE Code 10.20. This rule implies the exclusion of enterprises involved principally in other activities and has major implications for time trends in the size of the processing sector. For example, annual variations in income from processing in enterprises that conduct both processing and trade might lead to alternative annual inclusion and exclusion of these enterprises from the DCF estimates of total processing activity. STECF **recommends** that a study be conducted to evaluate the effect of this selection rule and to assess other possibilities to collect more accurate data on the processing sector.

STECF reviewed comments made in the report of the STECF-SGECA 10-03 WG on the results of the call for economic data launched to produce the draft report on the "Economic Performance of the fish processing: Annual Report 2009". STECF notes that several MS failed to comply with the data upload deadline and did not send all the required data. STECF **recommends** all MS submit data by the deadline and thoroughly check the data first. In particular, MS should compare the data with available EUROSTAT data, to ensure quality and homogeneity in official sources. In case of discrepancies, STECF **recommends** MS should explain in the National Annual Reports, the reasons for such discrepancies.

STECF observes that excluding the variable "raw materials per species" from the DCF affects the general usefulness of including the processing sector in the DCF because the link with the catching sector cannot be evaluated. For the STECF opinion, see section 4.8 of this plenary report, reviewing the report of the STECF-SGECA 10-04 WG which produced the processing sector AER.

STECF, taking into account proposals from Regional Coordination Meetings (RCM) and the report of SGECA 10-03, agrees with the need to hold the following three workshops and understands that they are already planned by the Commission:

- Workshop on **calculation of capital value** in accordance to PIM methodology and definition of variables not clearly defined in the DCF
- Workshop on **statistical issues** related to the collection of economic data within the DCF
- Workshop on **allocation of Economic Data at disaggregated level** (e.g. metiér)

STECF suggests the following terms of reference for the three workshops.

*Workshop on **calculating capital value** using PIM and definition of DCF variables*

1. Present and discuss MS experiences in approaches and results estimating fleet capital value.
2. Hold a training session on the application of the Perpetual Inventory Method
3. Compare price per capacity unit applied by different MS and assumptions made on the PIM method (age schedules, depreciation schemes, depreciation rates, etc.)
4. Discuss method problems faced by MS with respect to estimating unpaid labour and financial position, and propose solutions to the problems.
5. Propose clear definitions of those variables not clearly defined in the DCF.
6. Propose best practices to be followed by MS in estimating capital value using PIM.

*Workshop on **statistical issues** related to the collection of economic data within the DCF*

1. Present national methods to define sample size, accuracy indicators and estimate results.
2. Identify best practices regarding estimation procedures, assessing quality of data collected and define minimum targets for quality of economic data.
3. Evaluate methods, advantages and disadvantages of collecting data using non probability sampling surveys. Consider the results of the proposed Study to Standardise Quality Reporting and Propose Methods in the case of Non-probability Sample Survey.

4. Address Non-Response issues, including how Non Response can influence quality. Propose methods to deal with high level of non response.
5. Prepare Guidelines to MS for best practices in statistical analysis and on how to define and select the appropriate sample sizes to be proposed in National Programmes.

*Workshop on **allocation of Economic Data at disaggregated level** (e.g. metiér)*

1. Identify needs of applications, e.g. Long Term Management Plans, Regional Analyses for funding purposes and Ecosystem Approach to Fisheries Management.
2. Identify methods to allocate earnings and costs (operating costs, labour costs, capital costs) at different aggregation levels. Consider the identification of cost drivers. Transversal variables could serve for this purpose. Consider vessels that may be active in more than one fishing metiér during the same year.
3. Propose a method to split economic variables among different areas when appropriate.
4. Assess data quality requirements of allocation methods with regard to particular characteristics of DCF data sources at each MS (e.g. logbooks).

ANNEX 1 THE REPORT OF THE SUB GROUP ON ECONOMIC AFFAIRS (SGECA 10-03). REVIEW OF ECONOMIC DATA COLLECTED IN RELATION TO THE DCF, HARMONISATION OF SAMPLING STRATEGIES

1. EXECUTIVE SUMMARY AND RECOMMENDATIONS

The objectives of the SGECA 10-03 meeting were to review the methodologies for economic data collection, to advice on calculation and presentation of quality indicators, to evaluate the situation regarding the response by MS to the call for economic data of the processing industry and to propose TORs for workshops in the field of economic data collection.

SGECA 10-03 evaluated NP proposals with the specific objectives to compare different approaches and to discuss on possible harmonisation of sampling strategies among MS. The group pointed out several methodological issues that need particular attention and that require specific advice. With regard to the survey methods, SGECA observed that most MS apply simple or stratified random sampling and considered useful to provide guidelines to illustrate statistical techniques and best practices.

SGECA 10-03 addressed the issue of high non response rate that is expected from census and from sample surveys. SGECA 10-03 recommended MS to increase the effort in acquiring responses from the sample or from the population, as the level of non response is affected by the methods used to carry out the survey. SGECA 10-03 also recalled the recommendation made by SGECA 09-02 to launch a call for a study to address the question of the impact of non random non response on the final estimates.

The group discussed the problem of coherence and comparability of economic variables when derived from different sources (e.g. surveys, fleet register, logbooks, sales notes). SGECA 10-03 recommended MS to cross-checking data coming from different sources (for instance landings coming from logbook and from surveys) and to use proposed methods to assess the coherence of estimates of different variables coming from different data sources.

With regard to the processing sector, SGECA 10-03 observed that official statistics on economic data on the processing sector are already available. Nevertheless, several MS carry out specific surveys to estimate economic data required by the DCF. SGECA 10-03 considered that, according to statistical best practices, duplication of data collection should be avoided. In case official statistics cannot be used to meet the requirements of DCF, MS should clearly explain the reason and justify the use of additional surveys.

SGECA 10-03 suggested guidelines on how MS should collect and present information on quality of the data collected. These guidelines include proper definitions of accuracy indicators to be presented by MS in the AR and type of presentation for each indicator in order to ensure the possibility to evaluate and compare the results.

SGECA 10-03 discussed on the procedure to be applied by STECF in order to evaluate the quality and agreed that specific targets (in terms of precision level, acceptable response rates, etc.) are necessary in order to assess the quality of data collected. Definition of precision targets is a difficult exercise considering the large number of different methodologies applied by MS, the high variability of each variable and the great number of fleet segments. To overcome these difficulties,

the group suggested a step by step process. In a first step (that is for 2011 AR), targets for quality assessment will be qualitative ones: transparent and complete description of methodologies, analysis of accuracy indicators reported by MS in the AR. The next step will have to be performed by a working group on statistical issues that will analyse and evaluate quality indicators. This working group could be convened annually and will be attended by national experts. In the medium term, this analysis will lead to the definition of some targets.

SGECA 10-03 discussed some methodological issues related to calculation of capital values. The WG observed that the application of the PIM methodology as required by the DCF represents a first step and a great advance in the fishery data collection for EU MS and it grants, by mean of a standard procedure, to obtain a good comparability of data on the capital value of fishing fleets among MS. In addition to methodological issues, SGECA 10-03 discussed some conceptual issues and experiences gained from the application of the PIM approach as related to the determination of capital values and costs derived. The WG observed that the application of the PIM methodology allows a highly standardised calculation. However, the wide range of possible input data (interest rates, depreciation times and methods) and the uncertainties involved requires further steps of clarification and harmonisation.

The group also considered that the use of book value in order to estimate capital value and capital costs will limit the use of data to a fiscal accounting analysis more than to an economic valuation.

SGECA 10-03 discussed and advised on the following issues: use of historical or replacement value, type of depreciation scheme and interest rates to be used for opportunity cost calculation. However, SGECA considered that there are other issues still open that would need clarification and also considered that several MS are still facing technical and methodological problems to apply the PIM methodology. Therefore, SGECA recommended to held a workshop on capital value and proposed specific TORs for this meeting.

The Group revised the comments made by JRC on the results of the call for economic data launched to produce the draft report on the "Economic Performance of the fish processing: Annual Report 2009". SGECA 10-03 observed that several MS failed to comply with the data upload deadline and did not send all the required data. SGECA 10-03 recommended all MS to submit data in the given time frame and thoroughly check the data quality before submitting them.

JRC also indentified important divergences for some MS on the data submitted by MS for the elaboration of the processing industry report and the EUROSTAT data for fish processing (NACE 10.2). SGECA 10-03 considered that MS should check the data, mainly with the available EUROSTAT data, to ensure quality and homogeneity in official sources. In case of divergences, SGECA 10-03 recommended that MS should explain and justify in the AR the reasons for these divergences.

SGECA 10-0 3 also observed that the exclusion of the variable “raw materials per species” from the new DCF affects the general usefulness of the inclusion of the processing sectors in the DCF because the link with the catch sector cannot be evaluated anymore. SGECA 10-03 recognized the complexity of the issue and recommended to launch a study to investigate this point, in the case a concrete interest in the linkage between the processing industry and the catching sector exists. The study should also analyse the feasibility to collect data and eventually indicate appropriate methodologies.

SGECA 10-03, taking into account RCM proposals and outcomes of the present meeting, agreed on the necessity to held the following 3 workshops:

- Workshop on calculation of capital value in accordance to PIM methodology and definition of variables not clearly defined in the DCF

- Workshop on statistical issues related to the collection of economic data within the DCF

- Workshop on allocation of Economic Data on disaggregated level (e.g. metier)

SGECA 10-03 proposed TORs for these workshops.

Finally, DGMARE asked SGECA 10-03 to provide a proper definition of the processing industry/activities. SGECA 10-03 observed that the Statistical classification of economic activities in the European Community (NACE Rev. 2) includes a very detailed description of activity 10.20 "Processing and preserving of fish, crustaceans and molluscs" and recommended MS to refer to this description when identifying the population that should be surveyed.

2. INTRODUCTION

SGECA 10-03 met in Salerno, Italy from 20th to 24th September 2010. The terms of reference for the meeting are given below.

This meeting was a follow up of SGECA 09-02, that recommended the inclusion of a methodological report in the NP proposals and that suggested accuracy indicators for assessment of quality of estimates.

In addressing its TORs, SGECA 10-03 also took into account recommendations from Regional Co-ordination Meetings and from SGRN 10-01 and 10-02.

2.1. Terms of Reference

The specific terms of reference for SGECA-10-03 were as follows:

2. Evaluation of the methodological reports included in NP proposals for the years 2011-2013 in order to compare different approaches. Discussion on possible harmonisation of sampling strategies among MS

3. Analysis of accuracy indicators achieved by MS in order to verify the possibility to define specific precision targets or sampling rates for each fleet segment and economic variable listed in appendix VI of the DCF

4. Discussion of methodological issues related to calculation of capital values and clarification of methods and definitions.

5. Response by MS to the call for economic data launched to produce the draft report on the "Economic Performance of the fish processing: Annual Report 2009 ": To evaluate the situation regarding the response by MS to the call for economic data launched to produce the draft report on the "Economic Performance of EU fish processing: Annual Report 2009 " on both coverage and quality of the data submitted. Data failures will be identified by the group in order to allow the Commission to enforce MS obligations on a clear basis. It is expected that quality, comparability and coherence issues will be raised by the economists, with particular emphasis on quality checks.

6. Definition of TORs, content and objectives of proposed workshops for 2011 on economic data

7. AOB.

2.2. Participants

The full list of participants at SGECA 10-03 is presented in the following table:

Name	Address	Telephone no.	Email
STECF members			
Sabatella, Evelina (Chair)	IREPA onlus Via San Leonardo, trav. Migliaro 84131 Salerno Italy	Tel: +39 089 338978 Fax: +39 089 330835	esabatellarepa.org
Fabienne Daures	IFREMER Centre de Brest - DEM-BP 70, 29280 BREST France		fdaures@ifremer.fr
Aaron Hatcher	University of Portsmouth Winston Churchill Avenue PO1 2UP Portsmouth United Kingdom		aaron.hatcher@port.ac.uk
External experts			
Heleen Bartelings	LEI, Burgemeester Patijnlaan 19, 2585 BE The Hague Netherlands	+ 31-(0)70-3358350	heleen.bartelings@wur.nl
Rickard Bengtsberg	Swedish Board of Fisheries Ekelundsgatan 1, PO Box 423, 40126 Göteborg Sweden		rickard.bengtsberg@fiskeriverket.se
Jörg Berkenhagen	VTI-Federal Research Institute for Rural Areas, Palmaille 9, 22767 Hamburg Germany	+49 040 38905-206	joerg.berkenhagen@vti.bund.de
Irina Davidjuka	Fish Resources Research Department Daugavgrivas 8 LV-1048 Riga Latvia	+37 167 617 527	irina.davidjuka@bior.gov.lv
Loretta Malvarosa	Irepa Onlus Via S. Leonardo 84131 Salerno Italy		malvarosa@irepa.org
Michele De Meo	Irepa Onlus Via S. Leonardo 84131 Salerno Italy	+39089338978	DEMEO@IREPA.ORG
Michael Ebeling	VTI-Federal Research Institute for Rural Areas, Palmaille 9, 22767 Hamburg Germany	040-38905212	Michael.Ebeling@ish.bfa-fisch.de

Leyre Goti	General Concha 44, 4ºcentro izda 48012 Bilbao Spain	+34675621801	leyregoti@yahoo.com
Arina Motova	European Regional Policy Institute S. Konarskio str. 49 LT-03123 Vilnius Lithuania	+37061219519	arina.motova@erpi.lt
Carlos Moura	DGPA - Direcção Geral das Pescas e Aquicultura Av. de Brasília 1449-030 Lisboa Portugal	+351966702075	cmoura@dgpa.min-agricultura.pt
Heidi Pokki	Finnish Game and Fisheries Research Institute Viikinkaari 4 00790 Helsinki Finland	358404871305	heidi.pokki@rktl.fi
Darcelle Vassallo	Agriculture and Fisheries Regulation Department Fort San Lucjan, BBG 1283 Marsaxlokk Malta	+35 622 293 315	darcelle.vassallo@gov.mt
JRC experts			
Jordi Guillen	Joint Research Centre Via E. Fermi, 2748, I-21026 Ispra, Italy	+39 0332 78 5383	jordi.guillen@jrc.ec.europa.eu
Jarno Virtanen	Joint Research Centre Via E. Fermi, 2748, I-21026 Ispra, Italy	+39 0332 78 9614	jarno.virtanen@jrc.ec.europa.eu
European Commission			
Angel Calvo	European Commission DG Fish, J-99		angel-andres.calvo-santos@ec.europa.eu
Jordi Guillen	Joint Research Centre Via E. Fermi, 2748, I-21026 Ispra, Italy	+39 0332 78 5383	jordi.guillen@jrc.ec.europa.eu
Jarno Virtanen	Joint Research Centre Via E. Fermi, 2748, I-21026 Ispra, Italy	+39 0332 78 9614	jarno.virtanen@jrc.ec.europa.eu

3. EVALUATION OF THE METHODOLOGICAL REPORTS INCLUDED IN NP PROPOSALS FOR THE YEARS 2011-2013 IN ORDER TO COMPARE DIFFERENT APPROACHES. DISCUSSION ON POSSIBLE HARMONISATION OF SAMPLING STRATEGIES AMONG MS

NP proposals 2011-2013 have been submitted by Member States under the new guidelines (SGRN 09-03) that take into account SGECA 09-02 recommendations.

In particular, SGECA-09-02 suggested that in order to obtain methodology descriptions of a comparable standard among Member States, a methodological report should be included in the national programmes. The aim of the methodological report is to provide a clear and detailed description of the data collection methodologies and it should include information on:

• how the data will be collected and compiled;

• details of sampling techniques, estimation methods and data sources;

• a description of the methods used to evaluate the quality of the estimates

SGECA 10-03 has been asked to evaluate NP proposals with the specific objectives to compare different approaches and to discuss on possible harmonisation of sampling strategies among MS.

The group discussed on the possible outcomes of this evaluation and agreed that NP proposals should be reviewed focussing on the methodological issues and on the strengths and weaknesses of different data collection approaches. The compliance of NP Proposals with DCF (requirements and guidelines) had not been evaluated, as this task is under the remit of SGRN.

The group also considered that the evaluation should cover the fishing sector as well as the processing and the aquaculture sectors. However, considering the time constraints and the range of expertise attending the meeting, it was decided to leave out the evaluation of methodologies for the collection of aquaculture data. However, SGECA 10-03 considered that a review of these methodologies could be highly relevant and could give valuable advice to MS, also considering that the collection of aquaculture data has been included in NP proposals only from 2009.

The group discussed on the best approach to carry out the evaluation of the methodological reports and agreed on the following procedure:

- develop a template. This template is composed by a set of specific questions divided into sections which are in line with the structure of the methodological reports. The template for the evaluation of the fishing sector is given in annex I and the template for evaluation of the processing industry is given in annex II
- evaluation of NP using the template. Experts have been divided into 3 groups, 2 dealing with the fishing sector and one dealing with the processing sector
- summary of the results of the evaluation into one unique table. These summaries are presented in annexes I and II for the fishing sector and the processing sector respectively
- discussion of the results of the evaluation. Final comments and recommendations are reported in the following paragraphs.

3.1. Evaluation of methodologies for the fishing sector

SGECA-10-03 observed that MS commonly use the same type of data collection for small scale fisheries and for other fleet segments. However, this does not imply that economic data for smaller vessels are easily accessible and that the quality of estimations is good. Evaluation of quality and completeness of information should be reported in the AR, following the guidelines reported in paragraph 4.

SGECA-10-03 observed that the use of fix panel is in most cases explained by a pragmatic approach in order to achieve a sufficient coverage rate. The group observed that fixed panel may or may not be a Non Probability Sample depending on whether the units within the panel have been randomly or not randomly selected.

The evaluation found out that for at least 5 fleet segments data cannot be provided for confidentiality reasons. These segments appear to be important ones (pelagic fleets over 40 m). SGECA 10-03 observed that this will lead to a substantial loss of information for certain fisheries and for national totals in some cases.

According to DCF, an inactive vessel is defined as a "vessel that has not been engaged in fishing operations during a year. For inactive vessels only capital value (Appendix VI), fleet (Appendix VI) and capacity (Appendix VIII) shall be collected". Some MS still apply different definition of "inactive vessels" (for instance based on analysis of main activity). SGECA 10-03 recommended to follow the regulation.

The group also observed that the identification of inactive vessels is an easy exercise when vessels are required to fill declarations of activities, as logbook. When these declarations are not available other methods are used, such as estimation of activity level from sales notes or specific sample survey.

SGECA 10-03 observed that almost all MS used the Community Fishing Fleet Register as the frame to access to population units. This is in line with the DCF which defines the population as all vessels in the Community Fishing Fleet Register. There are still cases where the frame is constituted by the list of vessels "commercial active" or with a level of income higher than a fixed threshold. In these very few cases, MS are required to describe how data are estimated for the whole population.

The group discussed the problem of coherence and comparability of economic variables when derived from different sources (e.g. surveys, fleet register, logbooks, sales notes). The use of different data sources is a common best practise that however requires an assessment of the extent to which the data from these different statistical processes can reliably be combined. SGECA 10-03 encouraged MS to cross-checking data coming from different sources (for instance landings coming from logbook and from surveys) and to use appropriate methods to assess the coherence of estimates of different variables coming from different data sources. SGECA 10-03 also recalled the list of methods and references that could be used to check for coherence and comparability suggested by SGECA 09-02.

SGECA 10-03 observed that, even if a large number of survey methods exist, most MS apply simple or stratified random sampling. Therefore, SGECA 10-03 considered useful to provide guidelines to illustrate statistical techniques in case of Simple and stratified random sampling. These guidelines are given in annex III.

SGECA 10-03 observed that most MS determine the sample size on the basis of a fixed coverage rate and only in few cases on the basis of statistical methods (achievement of a target CV). The

group considered that sample size is an important element of the survey as it is closely related with the quality of final estimates. The group considered that an analysis of minimum planned rate for different fleet segments should be carried out in order to give guidelines on how to define the sample size. This analysis should be carried out in the workshop on statistical issues that is planned for 2011 (TORs for this workshop are reported in paragraph 6 of this report).

The group also evaluated the estimation procedures reported by MS in the NP. In most cases, the estimators used are not clearly described and only in few cases the estimator is clearly identified. SGECA 10-03 observed that most MS apply simple or stratified random sampling. Therefore, SGECA considered useful to provide guidelines with the identification of estimators to be used in the case of simple and stratified random sampling (see annex III).

SGECA 10-03 also observed that in case of non probability sampling survey estimation procedures (including non-responses), are not described in most cases. SGECA 10-03 considered that SGECA 09-02 recommended launching a study to harmonise quality reporting and propose methodologies in this specific situation. General terms of reference for this study are included in the report of SGECA 09-02 and have been endorsed by STECF (July 2009). SGECA 10-03 recommended to launch the study as soon as possible so to give the possibility to MS to consider the results of the study when presenting quality indicators in the 2011 technical report on activities performed in 2010 (to be delivered by May 2011).

Regarding the information on data quality reported by MS in their NP, SGECA 10-03 observed that in most cases, CV is reported as indicator of variability, in line with STECF/SGECA recommendations. However, several MS did not give any description of methodologies used to assess the quality.

Some MS also describe the models used to assess the quality of estimates in case of non probability sampling survey. SGECA 10-03 considered that the presentation of models in NP is a valuable information which does not require methodological evaluation.

The group also reviewed the problems in the collection of economic data faced by MS and described in the NP proposals. Most common problems relate to the estimation of capital value, capital cost and unpaid labour. SGECA 10-03 discussed the estimation of capital value and capital cost under TOR3.

Regarding imputed value of unpaid labour, the group considered that, even if no clear definition is given in the current DCF, the main problem concerns methodological issues. Therefore, the group recommended including the discussion of this parameter in the next workshop on variable definition that is planned for 2011 (TORs for this workshop are reported in paragraph 6 of this report).

Some MS also mentioned methodological problems to estimate financial position for small vessels. Also in this case, SGECA 10-03 considered that a workshop with participation of national expert could give appropriate guidance.

Another problem reported by several MS concerns high non response rate that is expected from census and even from sample surveys. SGECA 10-03 observed that the level of non response is affected by the methods used to carry out the survey (for instance postal questionnaires could lead to higher non response rates compared to face to face interviewers). In these cases, SGECA 10-03 recommended MS to increase the effort in acquiring responses from the sample or from the population. SGECA 10-03 also recalled the recommendation made by SGECA 09-02 to launch a call for a study to address the question of the impact of non random non response on the final estimates.

3.2. Evaluation of methodologies for the processing industry

SGECA 10-03 observed that official statistics on economic data on the processing sector are already available. Nevertheless, several MS carry out specific surveys to estimate economic data required by the DCF.

SGECA 10-03 considered that, according to statistical best practices, duplication of data collection should be avoided. In case official statistics cannot be used to meet the requirements of DCF, MS should clearly explain the reason and justify the use of additional surveys.

SGECA 10-03 recommended MS to describe in detail how data are collected even when these data come from official statistics. In order to ensure transparency, NP should include all definitions and methodologies applied to collect economic data regardless the source used (Ministry, National Statistical Offices, etc.).

A relevant problem for most MS is the high level of non response in case of census. This problem affects the estimation procedures and the quality of final estimates. SGECA10-03 observed that this issue has the same implications as in the fishing sector. Therefore, SGECA 10-03 considered that the required study on the impact of non response on the final estimates could give useful advice also for the collection of economic data of the processing sector.

SGECA 10-03 observed that most MS faced difficulties in estimation of the following variables: net financial costs, net extraordinary costs and imputed value of unpaid labour.

SGECA 10-03 suggested that, in order to harmonize definition adopted by MS, the IV Council Directive 78/660/EEC could be used as reference. In particular, financial costs are defined in art. 23, item C.13 of the IV Council Directive 78/660/EEC, ("Interest payable and similar charges, with a separate indication of those concerning affiliated undertakings"), while extraordinary costs are defined in art. 23, item 17, ("Extraordinary charges").

Regarding the imputed value of unpaid labour, SGECA 10-03 recommended to include this item in the workshop on capital and other variables that will be held in 2011.

SGECA 10-03 also observed that the current DCF states that methodology for the estimation of FTE should be as discussed in the report of Study FISH/2005/14, '2006. Calculation of labour including full-time equivalent (FTE) in fisheries". But this reference is clearly a mistake, because this methodology cannot be applied in the processing sector as it is based on an estimation of total hours of activity per metier.

3.3. Conclusions

SGECA 10-03 discussed the utility to carry out a methodological revision of National Programs.

The overall impression was that this kind of evaluation is useful in order to highlight points that should be treated in more detail and to identify strengths and weaknesses of different approaches applied by MS. Therefore, SGECA 10-03 recommended that SGRN should look at methodological aspects of NP and not only to compliance with DCF (regulations and guidelines). However, SGECA 10-03 is fully aware that the workload for SGRN during evaluations hardly allows for any additional task.

SGECA 10-03 considered that the inclusion of the methodological report in the NP proposals has improved the transparency of methods used by MS and this will lead to an increase of the general quality of the data collection programs.

4. ANALYSIS OF ACCURACY INDICATORS ACHIEVED BY MS IN ORDER TO VERIFY THE POSSIBILITY TO DEFINE SPECIFIC PRECISION TARGETS OR SAMPLING RATES FOR EACH FLEET SEGMENT AND ECONOMIC VARIABLE LISTED IN APPENDIX VI OF THE DCF

Table III_B_3 (Type of data collection scheme) requires MS to indicate, for each variable and for each fleet segment, the type of data collection, the accuracy indicators and the value of the accuracy indicators. All information have to be provided in the NP proposals, except the value of the accuracy indicators that should be reported in the Annual Report.

The group reviewed tables III_B_3 sent by MS in their AR 2010. During the meeting, tables III_B_3 of all MS have been aggregated in one spreadsheet in order to compare accuracy indicators achieved by MS and analyse the results. This exercise raised several issues that can be summarized into the following points:

- The column "type of data collection scheme" could include only 3 possible entries (A for A - Census; B for Probability Sample Survey and C for Non-Probability Sample Survey). But several MS did not follow guidelines

- The column "type of error" is aimed at classifying the accuracy indicators into two categories: variability and bias. Also in this case several MS did not follow guidelines. However, SGECA 10-03 recognized that the use of the term "bias" in this table could not be appropriate. The term "bias" in the statistical theory is often related to the bias of an estimator (the difference between an estimator's expectation and the true value of the parameter being estimated), while in this context the term "bias" is linked to information on coverage and response rates.

- The column "accuracy indicator" reports several entries (precision level, sample error, CV, coverage rate, variability, etc.). This is not in line with guidelines. However, SGECA 10-03 considered that the footnote under table III_B_3 is completely confusing (for bias: response rates and/or coverage rates and/or representativeness of the sample (always required in case of low response rate (<70%)). For variability: CV in case of B and variability of estimates in case of C). SGECA also considered that precise definition should be given to the required accuracy indicators.

- The column "value of the accuracy indicator" should report the actual values for CV and coverage/response rates. Even if information on quality indicators are not mandatory to be provided under NP 2009-10, several MS voluntarily provided quality data. However, figures under this column are not easily readable because MS use different format to present the same type of information and sometimes values are not consistent with information given in other columns.

- Analysis of results is also difficult because the table allows MS to group fleet segments into one line.

As a results of this exercise, SGECA 10-03 considered that the present guidelines for AR highly improved the presentation of information on quality with respect to previous programs. However, guidelines are still not clear and lead to several inconsistencies.

In particular, the group agreed on the importance to provide guidelines on how MS should collect and present information on quality.

4.1. Guidelines on how MS should collect and present information on quality analysis

SGECA-09-02 already discussed indicators to assess accuracy of estimates and suggested which information should be included in the annual report for each sampling strategies (e.g. census, probability sampling, non-probability sampling).

However, SGECA 10-03 considered that further information are needed in order to facilitate the inclusion of these indicators in the AR. In particular, accuracy indicators have to be better defined to avoid misinterpretation by MS.

The group also agreed with the necessity to assure consistency with guidelines for AR presentation and with previous definitions (already used by MS in the DCF context).

SGECA 10-03 proposals are summarized in table no. 2.

SGECA 10-03 discussed the issue of assessment of quality in case of non-probability sampling. There is no solution readily available to estimate the precision of estimators based on non-random sampling. However, the group agreed to follow the pragmatic approach proposed by the European Statistical System (ESS) standard quality reporting documents (EUROSTAT 2009a and 2009b). ESS suggests to assume probability sampling even in the case of non probability sampling in order to be able to use the CV. The group is aware that the value of this measure is questionable and therefore considered that the proposed study to harmonise quality reporting and propose methodology in case of non-probability sampling is of high relevance.

Table 2: Definition and presentation of accuracy indicators to be presented by MS in the AR

Type of data collection	**Accuracy indicators**	**Definition and *presentation***
A: Census B: Probability Sample survey	Achieved sampling rate (1)	achieved sample no.(2)/frame population no. *In % terms*
C: Non-Probability Sample survey	Response rate	achieved sample no.(2)/planned sample no.(3) *In % terms*
A: Census	Response rate	achieved no.(4)/ frame population no. *In % terms*
B: Probability Sample survey C: Non-Probability Sample survey A: Census, in case of low response rate (<70%) (5)	Coefficient of Variation (CV)	$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$ where: $\hat{\sigma}(\hat{Y})$ is the estimate of the standard deviation of $\hat{Y}$ $\hat{Y}$ is the estimate of the total *In absolute term (0.2 instead of 20%)*

(1) In case of census, achieved sampling rate is equal to 100%
(2) Achieved sample no. is the number of respondents and not, for instance, the number of questionnaires actually sent
(3) Planned sample no. is the number of sample units actually included in the survey
(4) Achieved no. is the number of respondents
(5) CV is required also for census because a census with very low response rate (<70%) has to be treated as a Non-Probability Sample survey

SGECA 10-03 also reviewed the guidelines for presentation of AR in order to take into account the present recommendations on how MS should collect and present information on quality analysis.

This revision has been carried out taking into account recommendations made by SGRN 10-02 (Evaluation of 2009 Annual Reports related to the Data Collection Framework). In particular, SGRN recommended quality indicators and figures in Tab. III.B.3 to be presented in the same way as in Tab. III.F.1.

SGECA 10-03 considered that table III_B_3 should include one line for each variable and for each fleet segment.

SGECA 10-03 also suggested that MS only provide the specification of variables in table III.F.1 for variables which is currently presented in table III.B.3 and duplicated in table III.F.1. Transversal variables are collected under appendix VI and also under appendix VIII. Therefore, to avoid duplication of information, the group considered that transversal variables should be reported only in table III.F.1. These are Fleet variables, days at sea and production value per species variables.

Regarding capital costs and capital value, SGECA 10-03 suggested to exclude them from table III_B_3. These variables are to be estimated according to the proposed PIM methodology in the capital valuation report of study No FISH/2005/03. The data and estimation procedures should be explained in the text of NP proposals. The inclusion of these variables in table III_B_3 does not add any relevant information.

The group agreed to include a new column in table III_B_1 with national name of survey (not mandatory). MS should name the surveys / data sources in the table in line with that in the text in NP / AR. The benefit from this is to simplify for future evaluators when references to specific surveys are made in the text.

The proposed revision of guidelines (text section III.B Economic variables, tables III_B_1, III_B_3 and III_F_1) are reported in annex IV.

The group revised only guidelines for section on the fishing sector. However, SGECA recommended that guidelines for processing and aquaculture sectors should be revised following the same approach used for the fishing sector.

4.2. Procedures to be applied for future evaluation of quality figures in NP and AR

DCF requires STECF to evaluate the quality of the data collected by the Member States (article 7, EU Reg. 199/08). However, so far, it is not clear how quality figures should be evaluated in future NP and AR. In this context, SGRN 10-02 recommended to provide principles for the evaluation of quality indicators.

SGECA 10-03 deeply discussed this issue.

The WG considered that each MS has the commitment to produce statistics of good quality. However, the evaluation of quality of data collection based only on review of NP proposals and AR may be subjective and time consuming.

Therefore, SGECA 10-03 agreed that specific targets (in terms of precision level, acceptable response rates, etc.) are necessary in order to assess the quality of data collected.

Definition of precision targets is a difficult exercise considering the great number of different methodologies applied by MS, the high variability of each variable and the great number of fleet segments. To overcome these difficulties, the group suggested a step by step process.

In a first step (that is for 2011 AR), targets for quality assessment will be qualitative ones: transparent and complete description of methodologies, analysis of accuracy indicators reported by MS in the AR.

The following step will consist in a working group on statistical issues that will analyse and evaluate quality indicators. This working group could be convened annually and will be attended by national experts. In the medium term, this analysis will lead to the definition of some targets.

The group also considered that targets have to be identified for different types of data collection.

In case of probability sampling surveys, targets can be fixed by analyzing the interval of confidence linked to different precision levels. This exercise would be carried out by the above mentioned working group on statistical issues where national experts will bring national data in order to run statistical techniques (an example of these techniques have been presented during the meeting and is attached to the report, see Annex V).

In case of census and Non Probability Sampling, precision targets should be defined in terms of different indicators (coverage errors, minimum level of responses,…) The suggested working group on statistical issues should also address this point.

SGECA 10-03 proposed the TORs for the working group on statistical issues (see paragraph 6, under TOR 4).

3. DISCUSSION OF METHODOLOGICAL ISSUES RELATED TO CALCULATION OF CAPITAL VALUES AND CLARIFICATION OF METHODS AND DEFINITIONS

A presentation was given to illustrate the PIM methodology that, according to DCF, MS have to apply to estimate capital value and capital cost.

The PIM has become one of the most important international standard for evaluation of tangible capital goods. "The perpetual inventory method (PIM) is the most widely used approach towards measuring stocks and flows of fixed assets. It rests on the simple idea that stocks constitute cumulated flows of investment, corrected for retirement and efficiency loss" (OECD manual, 2009).

The PIM generates an estimate of the capital stock by accumulating past purchases of assets over their estimated service lives. The application of the PIM to the fishery sector (fixed tangible assets – sea-based) can be summarised into the following steps:

- specification of the composition of the active fleet by age (vintage classes);
- estimation of the price per capacity unit (price/CU, e.g. per GT);
- the relative composition of the fleet in terms of the different assets (hull, engine, electronics and other equipment);
- calculation of the value of each vintage (year of construction) of the fleet applying the price/CU;
- calculation of capital costs by applying the most appropriate depreciation rates to each vintage's different assets and depreciation scheme (digressive or linear).

SGECA 10-03 observed that this methodology represents a first step and a great advance in the fishery data collection for EU MS. It grants, by mean of a standard procedure, to obtain a good comparability of data on the capital value of fishing fleets among MS.

Additionally, some experts discussed on the possibility to use book value in order to estimate capital value and capital costs. This approach would not require any estimation as data are already available in balance sheets. One expert also reported that, with regard to his national fisheries, outcome of the PIM method has led to an increase of the capital value by a factor of 3 to 4 as compared with previous results.

The WG considered that in any case balance sheet items can be used as input values for application of the PIM methodology that allows the estimation of both historical and replacement capital values.

There is, indeed, no one unique single method to evaluate the capital value and capital costs. The method depends on the analytical purpose. Two fundamentally different types of analysis are distinguished:

- Macro (economic) approach, which values capital at replacement (current) prices and accounts for opportunity costs. The analytical scope is focused on the sector, the society at large.

- Micro (fiscal) approach, which is close to fiscal accounting, values capital at historical prices and accounts only for interest costs paid. In this case the analytical scope is the individual firm level.

DCF requires MS to estimate both replacement and historical values. And this could create ambiguity when one value has to be chosen for economic analysis.

SGECA 10-03 considered that historical valuation implies that different vintages cannot be aggregated because each is on a different price basis. Valuation at historical prices is the usual procedure in company accounts. The disadvantages of using historical prices are due to the fact that assets which have been acquired at different dates are being valued at different prices so that when prices are rising/falling assets acquired more recently are implicitly given a higher/lower weight than those acquired in earlier periods. Capital stocks valued at historical prices cannot be compared with national accounting or other economic statistics that are expressed at prices of a single period (OECD manual, 2009)

Another methodological problem that has been raised by the application of the PIM method is the type of depreciation scheme (linear or geometric) to be applied. In the linear approach the value loss of the asset between two consecutive vintages is a constant amount. In the geometric or degressive approach, the value loss of the asset between two consecutive vintages follows an esponential trend (digressive approach).

SGECA 10-03 observed that OECD manual 2009 seems to be in favour of the geometric approach ("it has been used in a large number of economic studies and is also gradually adopted by statistical agencies").

Finally, SGECA 10-03 discussed on which interest rates should be used in order to calculate capital cost (real or nominal interest rates). The group did not reach any consensus and recommended to include this issue in the TORs of the next STECF plenary meeting.

SGECA 10-03 considered that there are other issues still open that would need clarification (for instance which depreciation rates MS should apply, national rates or harmonised (at EU level) rates).

At the same time, it seems that several MS faced technical and methodological problems to apply the PIM methodology. Therefore, SGECA recommended to held a workshop on capital value that should have the following objectives:

- train experts on the common tool developed for the application of the PIM methodology;

- compare assumptions made in the model (for instance service life for each type of asset, depreciation functions and depreciation rates, etc.).

TORs for this workshop are reported in paragraph 6 under TOR 5.

5. RESPONSE BY MS TO THE CALL FOR ECONOMIC DATA LAUNCHED TO PRODUCE THE DRAFT REPORT ON THE "ECONOMIC PERFORMANCE OF THE FISH PROCESSING: ANNUAL REPORT 2009 "

The Group revised the comments made by JRC on the results of the call for economic data launched to produce the draft report on the "Economic Performance of the fish processing: Annual Report 2009".

Data delivery time frames

DG MARE launched an official call for economic data of the processing sector on the 21st of September 2008, meaning that the deadline for the data submission was the 19th of October. The data call requested data for the years 2006 and 2007. From the 19th to the 23rd of October took place the SGECA 09-03 meeting in order to prepare the SGECA-09-03 report on the processing industry.

Some MS were still uploading corrected data long after the SGECA-09-03 meeting. This implied that the final version of the report was only made available at the end of the year.

SGECA 10-03 observed that the failure to comply with the data upload deadline affects the possibility to produce the report on the processing industry during the meeting. SGECA 10-03 further considered that the failure to comply with the deadline may have a serious effect on the working efficiency of the SGECA 09-03 meeting.

SGECA 10-03 recommended that more time is made available between the data upload deadline and the SGECA meeting to allow JRC to check the data and prepare the tables and templates for the experts attending the meeting.

General data coverage

Two countries did not submitted any data and MS that have submitted data provided partial data. There were a total of 17 parameters requested on the Fish Processing Sector Data Call. Only 61% of all data requested for 2007 was really submitted.

SGECA 10-03 recognized that missing data, delays in submission and divergences with other official data sources are MS responsibilities; SGECA recommended all MS to submit data in the given time frame and thoroughly check the data quality before submitting them.

In terms of data coverage, SGECA 10-03 observed that parameters more difficulties to be submitted have been raw materials per species, products price and capacity utilization. SGECA observed that these parameters have been excluded from the DCF wit the 2008 revision, because of the difficulties in their estimation. SGECA considered that quality of these data is also difficult to assess, therefore, it recommended not to use them in the analysis.

SGECA 10-03 also observed that the exclusion of the variable “raw materials per species” affects the general utility of the inclusion of the processing sectors in the DCF because the link with the catch sector cannot be evaluated anymore. However, SGECA 10-03 was aware of the complexity of the issue that should take into account also the origin of raw materials (imported or not, marine species or inland, etc.).

SGECA 10-03 recognized the complexity of the issue that is not reflected in the current DCF. If a concrete interest in the linkage between the processing industry and the catching sector exists, then SGECA 10-03 recommends to launch a study to investigate this point. The study should also analyse the feasibility to collect data and eventually indicate appropriate methodologies.

Data quality

SGECA 10-03 discussed the methods used by JRC to check for quality of data submitted.

JRC performed a series of manual checks on the data that basically consist in verifying the consistencies among related variables and in displaying unreliable and unlikely estimates.

Considering that it is the responsibility of the MS to deliver high quality data, the group suggested that MS should implement some routinely checks to assure minimum levels of quality before delivering data to the Commission.

JRC also indentified important divergences for some MS on the data submitted by MS for the elaboration of the processing industry report and the EUROSTAT data for fish processing (NACE 10.2). The indicators used to find out divergences have been employment and number of firms.

SGECA 10-03 considered that MS should check the data, mainly with the available EUROSTAT data, to ensure quality and homogeneity in official sources.

In case of divergences, SGECA 10-03 recommended that MS should explain and justify in the AR the reasons for these divergences.

6. DEFINITION OF TORS, CONTENT AND OBJECTIVES OF PROPOSED WORKSHOPS FOR 2011 ON ECONOMIC DATA

DGMARE informed that the last liaison meeting (Oostende, 2010) approved the recommendations made by several RCMs to hold workshops on methodological issues for the collection of economic data. These workshops will be convened in 2011 and will be attended by national experts appointed by the National Correspondents. SGECA 10-03 was asked to define the TORs for these workshops.

SGECA 10-03, taking into account RCM proposals and outcomes of the present meeting, agreed on the necessity of the following 3 workshops:

- Workshop on calculation of capital value in accordance to PIM methodology and definition of variables not clearly defined in the DCF
- Workshop on statistical issues related to the collection of economic data within the DCF
- Workshop on allocation of Economic Data on disaggregated level (e.g. metier)

SGECA 10-03 considered that another workshop on methods used to calculate effort data could be useful. SGECA 10-03 observed that there is a potential duplication of work with

respect to effort data requested by biologists and by economists. The workshop should examine the calculations used and where different calculation methods are necessary, the differences should be highlighted. The TORs for this workshop on transversal variable should be defined with biologists and the workshop should be attended by biologists and economists.

SGECA 10-03 also discussed on the utility to held a workshop on clustering methodologies. The WG observed that SGECA 09-02 defined some guidelines on how to cluster segments in order to have a common approach. Therefore, SGECA 10-03 agreed that MS should follow these guidelines.

6.1. Terms of reference of the proposed workshops

Workshop on calculation of capital value in accordance to PIM methodology and definition of variables not clearly defined in the DCF

1. 1. Present and discuss MS experiences in the approaches and results used to determine capital value.
2. Hold a training session on the application of the PIM method
3. Compare price per capacity unit applied by different MS and assumptions made on the PIM method (age schedules, depreciation schemes, depreciation rates, etc.)
4. Discuss methodological problems faced by MS with respect to estimation of unpaid labour and financial position
5. Propose best practices to be followed by MS.

Workshop on statistical issues related to the collection of economic data within the DCF

1. Presentation of national methods used to sample and estimate results by using a common template for the presentation.

2. Identification of best practices regarding the results, the quality of estimators and the context.

3. Analysis of methods used to collect data on non probability sampling surveys. Evaluation of advantages and disadvantages of non probability surveys. Discussion of the results of the study to harmonize quality reporting and propose methodology in the case of non-probability sample survey

4. Addressing Non-Response issues. How Non Response can influence quality. Methods to deal with high level of non response

5. Guidelines for best practices

Workshop on allocation of Economic Data on disaggregated level (e.g. metier)

1 Identify needs of exemplary applications, like:
- ♦ Long Term Management Plan
- ♦ Regional Analysis for funding purposes
- ♦ Ecosystem Approach to Fisheries Management

2 Identify methods to allocate earnings and particularly costs (operative costs, labour costs, capital costs) at different aggregation levels. This could/will include the identification of cost drivers. Transversal variables could serve for this purpose. The case of vessels that may be active in more than one fishing metiér during the same year should be considered especially.
3 Suggest a methodology to split, if necessary, economic variables among different spatial areas.
4 Assess data quality requirements of allocation methods with regard to particular characteristics of DCF data sources at each MS (e.g. logbooks)

7. AOB.

DGMARE asked SGECA 10-03 to provide guidance to Cyprus to solve a methodological issue on the processing sector.

Cyprus states (AR, page 22): "Some problems may be encountered due to the fact that the definition of the processing industry/activities it is not explicitly defined in the Decision. For example glazing and re-packaging is not cleared whether they are defined as processing activities? The Cyprus National Statistical Service does not consider them as processing activities. This inconsistency creates some problems. The Working Group on Processing Industry and Aquaculture: Review of economic issues (Ispra 2006) did not clarify this issue. There was a recommendation from the 2006 RCM for the Mediterranean area for further guidelines on the definition of processing industry activities."

SGECA 10-03 observed that the Statistical classification of economic activities in the European Community (NACE Rev. 2) includes a very detailed description of activity 10.20 Processing and preserving of fish, crustaceans and molluscs.

According to this reference, this class includes:
- preparation and preservation of fish, crustaceans and molluscs: freezing, deep-freezing, drying, cooking, smoking, salting, immersing in brine, canning etc.
- production of fish, crustacean and mollusc products: fish fillets, roes, caviar, caviar substitutes etc.
- production of fishmeal for human consumption or animal feed
- production of meals and solubles from fish and other aquatic animals unfit for human consumption

This class also includes:
- activities of vessels engaged only in the processing and preserving of fish
- processing of seaweed

This class excludes:
- processing and preserving of fish on vessels engaged in fishing, see 03.11
- processing of whales on land or specialised vessels, see 10.11
- production of oils and fats from marine material, see 10.41
- manufacture of prepared frozen fish dishes, see 10.85
- manufacture of fish soups, see 10.89

SGECA 10-03 recommended MS to apply this classification to define the list of processing enterprises to be included in the collection of economic data.

8. REFERENCES

NACE Rev. 2 – Statistical classification of economic activities in the European Community. Luxembourg: Office for Official Publications of the European Communities, 2008 — 363 pp. — 21 x 29.7 cm, ISBN 978-92-79-04741-1, ISSN 1977-0375

EUROSTAT (2009a). ESS Handbook for Quality Reports. ISBN 978-92-79-07855-2

EUROSTAT (2009b). ESS Standard for Quality Reports. ISBN 978-92-79-07854-5

ANNEX I TEMPLATE FOR EVALUATION OF THE FISHING SECTOR AND SUMMARY OF THE EVALUATION

SECTION 1 - TYPE OF DATA COLLECTION

Which type of data collection is applied for vessels less < 12 meters?
7 MS apply a census (MS with small fleets)
11 MS apply a Probability Sample Survey
6 MS apply a Non-probability Sample Survey

Which type of data collection is applied for other vessels?
For most MS, no difference in the approach from vessels < 12 m and vessels > 12 m

Which variables are estimated using official data (i.e. value of landings, day at sea, ...)?
Capacity, value of landings and most effort data come from official data for all MS, except for one (only capacity)

In case of Non Probability Sample survey, what reason (if any) is given for not using probability sampling?
In case of Non Probability Sample survey, the reason for not using probability sampling is in most cases to manage response rates and coverage rate (Data is already available and coverage rate is sufficient, Combining PSS with NPSS in order to increase sample rate, To avoid non response)

SECTION 2 - POPULATION

In case of clustering, what reason (if any) is given for clustering segments?
Clustering is made for confidentiality reasons in all cases. Only for one MS there is also a statistical reason.

Are there fleet segments for which data cannot be provided, for confidentiality reason?
5 MS have segments for which data cannot be provided for confidentiality reasons. These segments are often important ones

How are inactive vessels identified?
Inactive vessels: several MS defines inactive vessels in a different way (no main activity, on fleet register but not licensed,..) with respect to DCF definition (no days of activity)
Methods to identify inactive vessels not reported

In case frame population is different from target population, how data are estimated for the target population?
When this occurs, the difference is due to inactive vessels. In only one case it is explained how imputation will be carried out (Imputation based on official data and sampling data)

SECTION 3 - DATA SOURCES

Which data sources are used for vessels less < 12 meters?
In most cases, data sources include questionnaires and "official" data (like coastal logbooks, sales notes, vessel register). One MS uses only questionnaire

Which statistical methods or extra information are used to check coherence when estimation is based on different data sources?
When this information is given, in most cases, MS declare to cross-check volume and value of landings between logbooks and questionnaires (UK also uses VMS data). One MS applies an interesting approach: panel and logbook data checked for days at sea and landings.

SECTION 4 - SAMPLING

Which type of sampling strategy is applied (simple random sampling, systematic sampling, sampling with PPS, multiple stage sampling, etc.) for each segment?
In case of Probability sampling, most MS apply a Stratified Random Sampling. 3 MS apply more complex random sampling (PPS, Hierarchichal multi-stage, Systematic random sampling)
In case of Non Probability sampling: Voluntary sample. Poor information given

Which targets have been used to determine the sample size and why these targets have been chosen?
Information not included in all NPs
4 MS defines sample size on the basis of statistical methods (achievement of a target CV).
Other 4 MS defines the sample size on the basis of a fixed coverage rate

Do MS encounter any problems and in case yes, which problems?
Most common problems relate to the estimation of: capital value, unpaid labour.
Large non-response for small-scale vessels is also reported by 1 MS. Problems related with high non response rate is reported by other 4 MS

SECTION 5 - ESTIMATION

Is the type of estimators used (for example, Horvitz-Thompson or Hansen-Hurwitz estimators) reported?
Very few answers. 5 MS use Horvitz-Thompson (4) or Hansen-Hurwitz (1) estimators. In other cases, the estimator used is not clear
Some MS refer to "Simply raising"

Are estimation procedures described (including non-responses), also in case of non probability sampling survey?
Estimation procedures not described in most cases

SECTION 6 - DATA QUALITY EVALUATION

For each type of data collection, how variability is assessed?
In most cases, CV is reported as indicator of variability.

However, several MS did not follow guidelines

And bias?
In most cases, coverage rates and response rates are reported as indicator of bias.
However, several MS did not follow guidelines

Is the use of models in case of non probability sampling reported?
Models in case of non probability sampling reported in 3 NP proposals (not clear which models are).

ANNEX II TEMPLATE FOR EVALUATION OF PROCESSING SECTOR AND SUMMARY OF THE EVALUATION

SECTION 1 - TYPE OF DATA COLLECTION

Are different types of data collection considered for different enterprises size classes?
16 countries "No", 5 yes

Which type of data collection is applied ?
7 countries use Census and PPS, depending on the variable. 8 countries apply Census only, 2 apply Census and NPPS, 2 PPS only, 2 NPSS only.

In case of Non Probability Sample survey, what reason (if any) is given for not using probability sampling?

- "There do not have a complete list of enterprises. No other clear reason is mentioned"
- None

SECTION 2 – POPULATION

How the target population is defined?

- According NACE code 10.20, formerly 15.20: 16 countries follow implicitly (by referring to Business register) or explicitly.
- Based on try and error and according to past responses: 1 country.
- No definition: 1 country
- Following veterinary register: 2 countries.
- Based on secondary sources availability: 1 country

Are there segments for which data is not collected? If not why?
No for 17 countries, 2 countries unclear, 2 countries, Yes for 2 countries (one has no firms > 250 employees, one excludes by territory).

Are there segments for which data is not provided? If not, why?
2 countries where no data are collected, 2 because of confidentiality reasons (one merge with smaller segment, 1 will maybe not report), 1 unclear, 1 will exclude segment, 15 provide data for all segments.

Which parameter is used to stratify the population (i.e. number of employees, prevalent activity,..)?
Employees/FTE 12/1, (thereof 2 also using territory), type of legal entity or firm type 1, unclear in 2 cases, not relevant in 5 cases.

SECTION 3 - DATA SOURCES

Do MS use data coming from National Statistical Offices?
13 countries do, 8 do not.

What other data sources are used?
14 send out an own questionnaire, interviews (face to face and by telephone) are made by 3 countries, 2 have no other information source, 4 use published annual account statements, 3 business register, one uses data from the national administration body responsible for subsidies.

Which statistical methods or extra information are used to check coherence when estimation is based on different data sources?

- 1 country: Interviews performed to enterprises selected in the same sampling frame as other econ variables.
- 3 Countries: Official company accounts (where published) are used to cross-checked the information gathered from the questionnaires.
- One country: Cross-checking with no further information
- One country: Cross-checking by using data from veterinary register
- One country: Cross-checking by special national register (SIRENE/France)
- 14 countries: No information is given

SECTION 4 –SAMPLING

Which type of sampling strategy is applied (simple random sampling, systematic sampling, sampling with PPS, multiple stage sampling, etc.) for each segment?

- Stratified random sampling: 3 countries
- Stratified non-random sampling: 1 country
- No sampling: 10 countries
- No information given: 1 country
- PPS: 1 country,
- Simple random sampling: 3 countries, one using Neyman allocation methodology
- Self selecting random sampling: 2 countries

Which targets have been used to determine the sample size and why these targets have been chosen?

- 15 countries report no target
- 1 country: 15% in the "≤ 10" category, 25% in the "11 – 49" category, and 30% in the "50 – 249
- 1 country: 100%
- 1 country mentions only "response rate".
- 1 country: Response rate of 80%, no justification
- 1 country: "No targets are presented, information on targeted sample sizes: The target sampling rates have been chosen in order to maximise the effective use of the resources available of the survey, with a view to ensuring the sample is as representative as possible"
- 1 country: 21% (for >50 employees) due to the character of the secondary data source

What is the range of the coverage rate in case of sample?

- 1 country: 31 firms or 20% of the population
- 1 country: 15%

- 1 country: 40-90
- 1 country: 15-100% of enterprises (15% for enterprises with less then 20 employees) >90% of the turnover and >90% of the employees
- 1 country: 60%
- 6 countries: Not given
- 10 country: Not applicable

Do MS foresee any problems and in case yes, which problems?

- 3 countries: Low response rate, because it is not compulsory to answer the survey
- 1 country: Voluntary, rotation
- 1 country: The only problem definition of the processing industry.
- 16 countries: No

SECTION 5 –ESTIMATION

Is the type of estimators used (for example, Horvitz-Thompson or Hansen-Hurwitz estimators) reported?

- 3 countries: Horvitz-Thompson
- 13 countries: No
- 1 country: "Data are transformed using the methods which are most appropriate in the circumstances given the sample size and the data gathered. For example, data may be transformed using multiple linear regressions and log linear multiple regressions. The method used is dependent upon which of the methods shows the most significance between the variables and the calculated variable."
- 1 country: Extrapolation of the sampling results through application of the appropriate extrapolation coefficients.
- 1 country: Extrapolation by weighted average
- 2 countries: Table is not informative enough-closer look to NP necessary

Are estimation procedures described (including non-responses), also in case of non probability sampling survey?

- 1 country: The initial weights assigned to the sample units will be adjusted based on data referring to respondents and non-respondent of the selected sample.
- 1 country: Yes, for non response they use average growth rate
- 1 country: Mean value imputation is used to estimate the value for the total population.
- 4 countries: Yes, but closer look into NP necessary as table gives no detailed information.
- 14 countries: Not applicable or no information given

SECTION 6 - DATA QUALITY EVALUATION

For each type of data collection, how variability is assessed?

- 15 countries: No information is given, mainly because they only apply "Census".
- 5 countries: CV
- 1 country: Variability will be assessed for all of the data collected by a non-probability sample survey by an indicator of precision.

And bias?

- 9 countries: Non-response rate
- 3 countries: Coverage rate
- 2 countries: Non-response rate and coverage rate
- 7 countries: Not specified or unclear

Is the use of models in case of non probability sampling reported? If yes, which models?

- No model or information on this topic is mentioned
- In case of census, how is the procedure to deal with low response rate described?
- 1 country: Efforts are made to include information from non-respondents – for example, for certain companies where published accounts are available, information is drawn from the accounts to provide additional data points for the estimates for these parameters
- 1 country: Estimation of missing values by weighted average (weighting by turnover)
- 1 country: 100% response rate is expected given the exhaustivity of data source (financial statements)
- 1 country: There is no estimate of the variable if the response rate is below 70% after crossing the 3 sources. The procedure to estimate non-responses (when the response rate is up to 70%) is also described.
- 1 country: Non-response rates are calculated according ESS Handbook for Quality Reports if the problem occurs
- 1 country: They use extrapolation
- 15 countries: No information

ANNEX III GUIDELINES ON STATISTICAL TECHNIQUES FOR SIMPLE AND STRATIFIED RANDOM SAMPLING

A number of methods exist that can be used to select a sample. In the next paragraphs the most common sample designs are described.

1 Simple random sampling

Simple random sampling is a method of selecting n units out of the N such that everyone of the ${}_NC_n$ distinct sample has an equal chance of being drawn. In practice a simple random sample is drawn unit by unit. The units in the population are numbered from 1 to N. A series of random numbers between 1 and N is then drawn, either by means of a computer programme that produces such table. Random samples are particularly useful when little is known about the target population.

When a number that has been drawn is removed from the population for all subsequent draws, the method is also called random sampling *without replacement*. Random sampling with replacement is entirely feasible: at any draw, all N members of the population are given an equal chance of being drawn, no matter how often they have already been drawn.

2 Stratified random sampling

In stratified sampling the population of N units is first divided into subpopulations of N_1, $N_2,\ldots,N_H$ units, respectively. These subpopulations are non-overlapping, and together they comprise the whole of the population, so that:

$$N_1 + N_2 + \ldots + N_H = N \qquad (6)$$

The subpopulations are called strata. To obtain the full benefit from stratification, the values of the N_h must be known. When the strata have been determined, a sample is drawn from each, the drawings being made independently in different strata. The sample sizes within the strata are denoted by n_1, n_2, ..., n_H, respectively. If a simple random sample is taken in each stratum, the whole procedure is described as stratified random sampling.

Given the above definition, we can state that stratified random samples take advantage of additional information of the fishery. The idea is to group vessels into (what are considered to be) relatively homogeneous groups.

Strata could be defined according to the fleet segment as defined in the DCF ("a group of vessels with the same length class (LOA) and predominant fishing gear during the year, according to the Appendix III").

The potential for sample bias is reduced through using a stratified random sample. Since the population has been stratified, a balanced sample can be chosen that is more likely to be

representative of the population than a purely random sample. Consequently, smaller samples can be selected than using a purely random sample

The theory of stratified sampling deals with the properties of the estimates from a stratified sample and with the best choice of the sample size n_h to obtain maximum precision.

3 The estimation of sample size and allocation across strata

In the planning of a sample survey, a stage is always reached at which a decision must be made about the size of the sample. The decision is important. Too large a sample implies a waste of resources, and too small a sample diminishes the utility of the results. The decision cannot always be made satisfactorily; often we do not possess enough information to be sure that our choice of sample size is the best one. Sampling theory provides a framework to solve these problems.

The principal steps involved in the choice of a sample size are as follows.

There must be some statement concerning what is expected of the sample. This statement usually is in terms of desired limits of error.

Some equation that connects n with the desired precision of the sample must be found. The equation will vary with the content of the statement of precision and with the kind of sampling that is contemplated. One of the advantages of probability sampling is that it enables this equation to be constructed.

This equation will contain, as parameters, certain unknown properties of the population (for instance the variability of the investigated phenomenon). These must be estimated in order to give specific results.

Finally, the chosen value of n must be appraised to see whether it is consistent with the resources available to take the sample. This demands an estimation of the cost, labour, time and material required to obtain the proposed size of sample.

In the case of simple random sampling, the formula for the definition of the sample size, n, is the following:

$$n \geq \frac{N^2 S^2}{\varepsilon^2 Y^2 + N S^2} \qquad (7)$$

where:

N is the number of units in the population,

ε is equal to $\theta/2$, where θ is the maximum error that is accepted for the final estimates,

S^2 is an estimate of $V(Y)$, total variance of the phenomenon in the population and is given by:

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2 \qquad (8)$$

with

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \qquad (9)$$

In the case of single stage stratified sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling, the formula to calculate the sample size, for a maximum error of 2θ, and with a probability equal to P=95%, is the following:

$$n \geq \frac{\left(\sum_{h=1}^{H} N_h S_h\right)^2}{\theta^2 + \sum_{h=1}^{H} N_h S_h^2} \qquad (10)$$

where, H is the total number of strata, N_h is the population size in the stratum h, S^2_h is the estimate of the variance in the stratum h.

Once the sample size has been obtained, the allocation among strata has to be defined. The allocation across strata can be assessed on the basis of two different methods:

the proportional criterion, in each stratum we have the same number of units:

$$n_h = n/H$$

the Neyman criterion, by which a variable percentage of elements is drawn from each stratum in order to minimise the value of the variance. In this case, the formula is:

$$n_h = n\frac{N_h S_h}{\sum_{h=1}^{H} N_h S_h} \qquad (11)$$

In order to apply the previous formulas n. 7, 10 and 11 a pre-estimate of the $S^2_{(h)}$ variances is required; in other words, the variances of the target variables of the survey must be known. For this purpose, the results of previous sampling survey can be used. In the case this information is

not available, a pilot study can be developed to have an indication of the variability of the investigated phenomenon.

3.1 Sample size with more than one item

In most surveys information is collected on more than one item. Sometimes the number of items is large. In particular, fisheries surveys are generally multivariate, that is, the variables investigated are more than one (revenues, costs, employment, etc.).

If a desired degree of precision is prescribed for each item, the calculations lead to a series of conflicting values of n, one for each item. It may happen that the *n*'s required are all reasonably close. If the largest of the *n*'s falls within the limits of the budget, this *n* is selected. More commonly, there is a sufficient variation among the *n*'s so that we are reluctant to choose the largest, either from budgetary considerations or because this will give an over-all standard of precision substantially higher than originally contemplated. In this event the desired standard of precision may be relaxed for certain of the items, in order to permit the use of a smaller value of *n*.

However, nowadays with progress in sampling theory and software technology some methods are available to menage the problem of the sample size with more than one item.

In particular, the Bethel method is the application of Neyman's method to the multivariate case. The approach used by this method is to transform the analysis into a linear programming model that allows the identification of the sample size and the allocation across strata, minimising the variances of all variables simultaneously (see also Bethel, 1989).

4 Estimation of parameters

In the previous chapters we described how sampling theory can be employed to obtain information about samples drawn at random from a known population. From a practical viewpoint, however, it is also very important to be able to infer information about a population by use of samples drawn from it. Such problems are dealt with in statistical inference, which uses principles of sampling theory.

One important problem of statistical inference is the estimation of values referred to the population (such as population totals, means, variances, etc.) from the corresponding sample data.

In this guideline, the Horvitz-Thompson estimator is reported, but a lot of other estimators exist that have been studied by the sampling theory.

The Horvitz-Thompson estimator is used to estimate the total values (for examples total catches, total revenues, total employment, total costs, and so on).

In the case of simple random sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling the formula is:

$$\hat{Y}_{HT} = \sum_{i}^{n} w_i y_i \qquad (12)$$

where $w_i = N/n$, and y_i are the observed values from the sample units.

In the case of a stratified random sampling, and in the hypothesis of extracting the sampling units with equal probability and without re-pooling, the formula is:

$$\hat{Y}_{HT,h} = \sum_{i=1}^{n_h} w_{hi} y_{hi} \qquad (13)$$

with:

n_h sample size in the stratum h;

$w_{hi} = N_h/n_h$:

y_{hi}, sample data of the unit i in the stratum h.

5 Sources of error in surveys

The sample theory assumes that some kind of probability sampling is used and that the observations y_i on the ith unit is the correct value for that unit. The error of estimate arises solely from the random sampling variation that is present when n of the units are measured instead of the complete population of N units.

These assumptions hold reasonably well in the simpler types of surveys in which the measuring devices are accurate and the quality of work is high. In complex survey, particularly when difficult problems of measurement are involved, the assumptions may be far from true. Three additional sources of error that may be present are as follows.

Failure to measure some of the units in the chosen sample. This may occur because of their refusal to answer the questions.

Errors of measurement on a unit. The measuring device may be biased or imprecise. The respondents may not possess accurate information or they may give biased answers

Errors introduced in editing, coding and tabulating the results.

These sources of error necessitate to develop methods for computing standard errors and confidence limits that remain valid when the other errors are present.

5.1 Sampling errors

The standard deviation is used to provide information about the relative distribution around the estimates. The level of confidence is related to the amount of variation around the estimates, so it is related to the standard deviation. Also, from the Central limit Theory, the larger the sample size, the greater the confidence in the estimate. Therefore the level of confidence is related also to the size of the sample.

The standard error is an indicator of the level of confidence in the estimate. In the case of stratified random sampling, the standard error in each strata is given by

$$\hat{\sigma}\left(\hat{Y}_{HT}\right) = s \cdot \sqrt{\frac{N^2}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \cdot \sqrt{\frac{N^2}{n}\left(1 - \frac{n}{N}\right)} \qquad (14)$$

where s is the standard deviation of the sample in each strata.

The Coefficient of Variation (CV) is given by

$$cv\left(\hat{Y}_{HT}\right) = \frac{\hat{\sigma}\left(\hat{Y}_{HT}\right)}{\hat{Y}} \qquad (15)$$

where C_h is the estimate of the variation coefficient.

Standard errors provide information about the confidence interval around the estimates. Assuming that the distribution is normally distributed around the estimate value, there is a 95 per cent probability that the true population value is within 1.96 standard errors of the sample value. That is, there is a 95 per cent probability that

$$\hat{Y}_{HT} - 1.96 \cdot \hat{\sigma}\left(\hat{Y}_{HT}\right) < Y < \hat{Y}_{HT} + 1.96 \cdot \hat{\sigma}\left(\hat{Y}_{HT}\right) \qquad (16)$$

5.2 Effects of non-response

Information can only be collected from those individuals selected willing to participate in the survey. However, a rejection may occur for a number of reasons, such as bad timing of the survey. A problem with rejection is that there is no guarantee that the individual who does not want to participate is the same as those who do participate. In some cases, there may be a correlation between rejection and the characteristics of the individual being surveyed. In such cases, exclusion of these individuals may result in a biased sample.

An advantage of the stratified approach is that another individual in the group can be selected to replace the reject. While this still may lead to some bias in the sample, the bias is not likely to

be as great as in a purely random sample. This is because the reject is being replaced by another individual with similar characteristics, and the assumption is made that any individuals within the group are representative of the group. The greater the degree of stratification, the more likely this assumption is to hold.

Another way to treat the non-response is to adjust the initial weights w_{hi} on the basis of data referring to the responses (r_h) and the non-responses (s_h) of the sample (n_h). The method consists in multiplying the initial weights (w_{hi}) by a factor (d_h) equivalent to:

$$d_h = \frac{r_h + s_h}{r_h} \qquad (18)$$

The hypothesis under this method is that a homogeneity of response exists within strata (Sarndal, Swensson & Wretman 1992). And in fact, the related estimator is called RHG (response homogeneity group).

REFERENCES

Bethel, J. (1989) Sample allocation in Multivariate Surveys. *Survey Methodology*, 15(1): 47-57.

Sabatella E., Ramon Franquesa*, Manual of fisheries sampling surveys: methodologies for estimations of socioeconomic indicators in the Mediterranean sea,* FAO/General Fisheries Commission for the Mediterranean, Studies and Reviews, No. 73 2003

Särndal C. E., Swensson B & Wretman J. (1992): Model assisted survey sampling. Springer-Verlag: New York.

ANNEX IV REVISED GUIDELINES FOR AR PRESENTATION

Revised text is given in red

III.B Economic variables

[Insert here supra-region header, according to Appendix II of Commission Decision 2008/949/EC. For each supra region, sections III.B.1-4 should be given.]

III.B.1 Achievements: Results and deviation from NP proposal

Update standard tables III.B.1, III.B.2 and III.B.3 with the information collected during the sampling year.

Description of fields in the table III.B.1: Population segments for collection of economic data

Fields	**Description/definition of the fields**
Achieved Sample no.	achieved sample no. should refer to the no. of respondents (and not for instance to the number of questionnaires actually sent) the no. of respondents should refer to the survey (unit response rate) and not to the variables (item response rate)
Achieved Sample rate	Achieved sample no./frame population no the no. of respondents should refer to the survey (unit response rate) and not to the variables (item response rate).

Description of fields in the table III.B.2: Economic Clustering of fleet segments

Fields	**Description/definition of the fields**
Total number of vessels in the cluster by the 1st of January of the sampling year	Updated number of vessels comprised in each of the clusters.
Number of vessels in the segment by the 1st of January of the sampling year	Updated number of vessels comprised in each of the fleet segments.

Description of fields in the table III.B.3: Economic Data collection strategy

Fields	**Description/definition of the fields**

Value of the accuracy indicators	Provide the value of the achieved accuracy as named in column I

In case of census with a very low achieved response rate (<70%), MS has to evaluate the representativeness of the data collected on the respondents.

List the shortfalls (if any) in the achieved data collection compared to what was planned in the relevant NP proposal, and explain the reasons for the shortfalls. Explain any deviation from the sampling intensity proposed, the methods used for collecting data and for estimating the parameters.

MS are reminded of the fact that the DCF has no provisions for the exclusion of any part of the vessel population from data collection (by means of thresholds for, e.g., fishing effort, quantities landed, revenues, etc.). If, nonetheless, part of the fleet was excluded from data collection, the reasons for this should be thoroughly explained and justified.

III.B.2 Data quality: Results and deviation from NP proposal

Update standard tables III.B.3 with the values of the accuracy indicators.

Accuracy indicators have to be reported for each fleet segment and for each variable.

Indicators of accuracy to be presented by MS in the TR are described in the following table:

Type of data collection	**Accuracy indicators**	**Definition and *presentation***
A: Census B: Probability Sample survey	Achieved sampling rate (1)	achieved sample no.(2)/frame population no. *In % terms*
C: Non-Probability Sample survey	Response rate	achieved sample no.(2)/planned sample no.(3) *In % terms*
A: Census	Response rate	achieved no.(4)/ frame population no. *In % terms*
B: Probability Sample survey C: Non-Probability Sample survey A: Census, in case of low response rate (<70%) (5)	Coefficient of Variation (CV)	$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$ where: $\hat{\sigma}(\hat{Y})$ is the estimate of the standard deviation of $\hat{Y}$ $\hat{Y}$ is the estimate of the total

		In absolute term (0.2 instead of 20%)
C: Non-Probability Sample survey A: Census, in case of low response rate (<70%)	Other variability indicators	In case of Non-Probability Sample survey, other measures of variability of estimators can be reported; description of these indicators should be reported in the text

Footnotes:

In case of census, achieved sampling rate is equal to 100%

Achieved sample no. is the number of respondents and not, for instance, the number of questionnaires actually sent

Planned sample no. is the number of sample units actually included in the survey

Achieved no. is the number of respondents

CV is required also for census because a census with very low response rate (<70%) has to be treated as a Non-Probability Sample survey

List the shortfalls (if any) in the achieved accuracy compared to what was planned in the relevant NP proposal, and explain the reasons for the shortfalls.

III.B.3 Follow-up of Regional and international recommendations

List the appropriate recommendations from all relevant RCMs related to the economic variables and give a brief description of the responsive actions taken. Use sub-headers to make the distinction between the different RCMs, and print recommendations and responses in a different font style (e.g. bold and/or italic for the recommendations and normal text for the descriptions of the action taken). There is no need to also list recommendations that do not apply to MS (e.g. on the terms of reference of ICES expert groups, on actions to be taken by the EC, etc.).

In doing so, you may have to go back several years in time and refer to RCM reports of more than one year ago. Most of the RCM recommendations and proposed actions will only take effect in the year following the actual meeting of the RCM and the actions taken by MS will only become visible in the Technical Reports that are submitted two or three years later.

III.B.4 Actions to avoid shortfalls

Briefly describe the actions that will be considered / have been taken to avoid the shortfalls in the future and when these actions are expected to produce effect. If there are no shortfalls, then this section can be skipped.

Table III.B.1 - Population segments for collection of economic data

										NP years		
										TR Year		
MS	**Supra region**	**Fleet segment (c)**	**Reference year**	**Target population no. (b) ----- N**	**Frame population no. (d) ---- F**	**Planned sample no. (a) (b) ----- P**	**Planned sample rate (a) ----- (P/F)*100 (%)**	**Type of data collection scheme**	**Achieved Sample no.**	**Achieved Sample rate**	**Achieved Sample no. / Planned sampled no.**	**National name of the survey (e)**
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Beam trawlers : 18-24 m*	2008	150	140	100	67	A			0,00%	
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Beam trawlers : > 40 m	2008	25	25	20	40-80	B				
ESP	Mediterranean	Passive gears : Drift and fixed nets 12-18 m	2008	5	5	5	100	C				
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Beam trawlers : 18-24 m*	2009	150	100	100	67	A				
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Beam trawlers : > 40 m	2009	25	25	15	40-80	C				
ESP	Mediterranean	Passive gears : Drift and fixed nets 12-18 m	2009	5	5	5	100	C				

(a) Where planned sample nos. and rates differ for the estimation of different parameters within a segment, please give the appropriate range.

(b) planned sample can be modified based on updated information on the total population (fleet register)

(c) put an asterisk in the case the segment has been clustered with other segment(s)

(d) For economic variables to be collected only for active vessels, the frame may be different from the population.

(e) name of the survey as reported in the NP if applicable. Not mandatory

Table III.B.3 - Economic Data collection strategy

										NP years	
										TR Year	
MS	**Supra region**	**Variable group (a)**	**Variables**	**Reference year**	**Data sources**	**Fleet segments (b)**	**Type of data collection scheme (c)**	**Achieved sample rate**	**Response rate**	**CV**	Other variability indicators (d)
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Income	Gross value of landings	2010	logbook	beam trawlers <6 m, beam trawlers 12-18 m	A	100%	85%		
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Income	Other income	2010	questionnaires	beam trawlers <6 m, beam trawlers 6-12 m	B	23%	90%	0,03	
ESP	Baltic Sea, North Sea and Eastern Arctic, and North Atlantic	Energy costs	Energy costs	2010	questionnaires	beam trawlers <6 m, beam trawlers 6-12 m	C	18%	60%	0,025	

(a) capital value, capital costs and transversal variables should not be reported in this table. Transversal variables have to be reported only in table III_F_1

(b) MS should specify the segments for which a specific sampling strategy has been used. CV, achieved sample rate and response rate have to be reported for each segment and each variable

(c) A - Census; B - Probability Sample Survey; C - Non-Probability Sample Survey

(d) only in case of Non Probability Sampling, measures of variability other than CV could be provided and explained in the text

Table III.F.1 – Transversal Variables Data collection strategy

NP years	
TR Year	

MS	Region	Variable group	Variables	Data sources	Fleet segments (a)	Type of data collection scheme (b)	Reference year	Achieved sample rate (c)	Response rate (c)	CV (c)
FRA	North Sea and Eastern Arctic	Capacity	Number of vessels			A		100%	100%	
FRA	North Sea and Eastern Arctic	Capacity	GT, kW, vessel age,							
FRA	North Sea and Eastern Arctic	Capacity	------							
FRA	North Sea and Eastern Arctic	Effort	Number of vessels							
FRA	North Sea and Eastern Arctic	Effort	Days at sea							
FRA	North Sea and Eastern Arctic	Effort	Hours fished							
FRA	North Sea and Eastern Arctic	Effort	Fishing days							
FRA	North Sea and Eastern Arctic	Effort	kW* fishing days							
FRA	North Sea and Eastern Arctic	Effort	-------							
FRA	North Sea and Eastern Arctic	Effort	-------							
FRA	North Sea and Eastern Arctic	Effort	--------							
FRA	North Sea and Eastern Arctic	Landings	Value of landings total and per species							
FRA	North Sea and Eastern Arctic	Landings	Live weight of landings total and per species							
FRA	North Sea and Eastern Arctic	Landings	-----							
FRA	North Sea and Eastern Arctic	Landings	-----							
FRA	Mediterranean and Black Sea	Capacity	Number of vessels							

FRA	Mediterranean and Black Sea	Capacity	GT, kW, vessel age,							
FRA	Mediterranean and Black Sea	Capacity	-------							
FRA	Mediterranean and Black Sea	Effort	Number of vessels							
FRA	Mediterranean and Black Sea	Effort	Days at sea							
FRA	Mediterranean and Black Sea	-------	--------							

(a) MS should specify the segments for which a specific sampling strategy has been used. CV, achieved sample rate and response rate have to be reported for each segment and each variable

(b) A - Census; B - Probability Sample Survey; C - Non-Probability Sample Survey

(c) to be reported only on a yearly basis

ANNEX V PRESENTATION ON STATISTICAL ISSUES

Coefficient of Variation and Confidence Interval

Michele De Meo

Irepa Onlus

IREPA

In order to provide some guidelines for calculating CV, it is necessary to introduce some simple concepts.

A **unit** is a single individual or object to be measured.

The **population** (or **universe**) is the entire collection of units about which we would like information.

IREPA

The first idea to obtain information should be to investigate the whole population.

Even if it may be more accurate than a sample, a census takes more time and resources to complete.

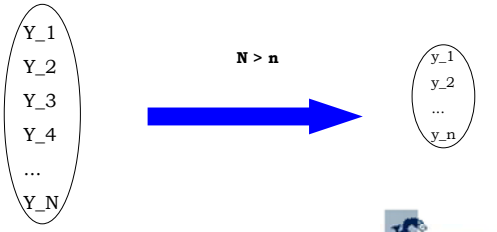
In this case, that is if a census is not appropiate, the only possibility is a **sample survey**

Some specification

We can simply define the **sample** as *a subgroup of the population*

While a **useful sample** is only a *sample representative of the universe* !!!!!!

From the population to the sample



For example, Y should be the fuel cost:

i = vessel code

Y_i = fuel cost for the vessel i

N = population size

n = sample size

We want to estimate $Y = \sum_{i=1}^{N} Y_i$

If we collect information with a simple random sampling, we estimate Y with $\hat{Y}$

$$\hat{Y} = N \cdot \bar{y} = \frac{N}{n} \sum_{i=1}^{n} y_i$$

An estimate of a quantity of interest, such as the *total fuel cost*, will generally be subject to sample-to-sample variation, that is the **sampling error** : the error caused by observing a sample instead of the whole population

To evaluate the quality of our data, we have to calculate the **coefficient of variation** of the estimate:

$$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$$

In our example:

$\hat{\sigma}(\hat{Y})$: the estimate of the standard deviation of $\hat{Y}$

$\hat{Y}$: the estimate of the total fuel cost $\hat{Y}$

But exactly how to calculate the numerator and denominator of this formula ?

$$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$$

The simplest case is that of simple random sampling

As seen in the previous slide , the Total Estimate is as follows:

$$\hat{Y} = \frac{N}{n}\sum_{i=1}^{n} y_i = N \cdot \bar{y}$$

And what about the numerator, that is the standard deviation ? It is important to make clarifications

$$cv(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$$

Normally we are used to calculate the standard deviation using the following formula (in Excel, this formula is available as **STDEV**):

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

In our case, the previous formula is not correct because it ignores the ***finite population correction***. So, the correct formula is:

$$\hat{\sigma}(\hat{Y}) = s \cdot \sqrt{\frac{N^2}{n}\left(1-\frac{n}{N}\right)} =$$

$$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2} \cdot \sqrt{\frac{N^2}{n}\left(1-\frac{n}{N}\right)}$$

and the previous formula can be easy calculated in Excel:

$$\hat{\sigma}(\hat{Y}) = STDEV \cdot \sqrt{\frac{N^2}{n}\left(1-\frac{n}{N}\right)}$$

Why the CV is so important?

the higher the level, the lower the accuracy

Than we can observe that the Coefficient of Variation is directly associated to the confidence interval

$$\hat{Y} \pm z_{\alpha/2} \cdot \hat{\sigma}(\hat{Y}) = \hat{Y} \pm z_{\alpha/2} \cdot \hat{Y} \cdot cv(\hat{Y})$$

The confidence interval increases with the coefficient of variation!

Example

Interval of confidence for the total fuel cost of 15000:

If **CV = 0.02**, than the **real** total fuel cost will be between

15000 - 1.96 * (15000*0.02) = **14412**

and

15000 + 1.96 * (15000*0.02) = **15588**

If **CV = 0.35**, than the **real** total fuel cost will be between

15000 - 1.96 * (15000*0.35) = **4710**

and

15000 + 1.96 * (15000*0.35) = **25290**

ANNEX VI: DECLARATIONS OF EXPERTS

Declarations of invited experts are published on the STECF web site on https://stecf.jrc.ec.europa.eu/home together with the final report.

European Commission

EUR 24635 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Scientific, Technical and Economic Committee for Fisheries. Report of the Working Group on review of economic data collected in relation to the DCF, harmonisation of sampling strategies

Author(s): Evelina Sabatella, Heleen Bartelings, Rickard Bengtsberg, Jörg Berkenhagen, Fabienne Daures, Irina Davidjuka, Michele De Meo, Michael Ebeling, Leyre Goti, Jordi Guillen, Aaron Hatcher, Loretta Malvarosa, Arina Motova, Carlos Moura, Heidi Pokki, Darcelle Vassallo, Jarno Virtanen

Luxembourg: Publications Office of the European Union
2010 – 60 pp. 21 x 29.7 cm
EUR – Scientific and Technical Research series – ISSN 1831-9424
ISBN 978-92-79-18752-0
doi:10.2788/58635

Abstract

This report reviews the methodologies for economic data collection and gives advice to presentation of quality indicators. The working group highlighted several methodological issues that need particular attention. SGECA 10-03 proposed procedures and guidelines on how to tackle these issues. Concerning the most difficult issues the working group suggested specific work shops that should produce clear practices on how to deal with these issues. Working group drafted a terms of reference for these meetings.

SGECA 10-03 gave recommendations on data collection on the processing industry. It noted that for some member states the data did not match that in EUROSTAT. The working group advised member states to explain the possible differenced with official data. The group also recognized the analytical importance of raw material use and suggested a feasibility study on collecting that data.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

The Scientific, Technical and Economic Committee for Fisheries (STECF) has been established by the European Commission. The STECF is being consulted at regular intervals on matters pertaining to the conservation and management of living aquatic resources, including biological, economic, environmental, social and technical considerations.

LB-NA-24635-EN-N

ISBN 978-92-79-18752-0